DOCUMENT RESUME

ED 048 337 TM 000 410

AUTHOR Wardrop, James L.
TITLE Determining "Most Probable" Causes: A Call for Re-Examining Evaluation Methodology.
PUB DATE Feb 71
NOTE 8p.; From symposium "Critique of the Report of the Phi Delta Kappa Study Committee on Evaluation," American Educational Research Association, New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Decision Making, Evaluation, *Evaluation Methods, *Evaluation Needs, *Evaluation Techniques, *Inductive Methods
IDENTIFIERS *Phi Delta Kappa Study Committee on Evaluation

ABSTRACT

Throughout the evaluation approach presented in this report, explanation (determination of the most probable cause or causes for a phenomenon) is crucial. It is argued that the Report of the Phi Delta Kappa Study Committee on Evaluation does not adequately treat this concept, nor does it adequately consider some implications of the concept for the methodology of evaluation. The centrality of explanation to evaluation should be emphasized. This perspective highlights existing evaluation methodology (context, input, process, and product) in terms of appropriateness to decisions to be served (explanations to be made) in the evaluation process. This approach identifies gaps in current methodology and provides clues as to how to close those gaps. (Author/DG)

ED 048337

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

DETERMINING "MOST PROBABLE" CAUSES: A CALL FOR RE-EXAMINING EVALUATION METHODOLOGY*

James L. Wardrop
Center for Instructional Research and Curriculum Evaluation
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

TM 000 410

*Paper read at the symposium, "Critique of the Report of the Phi Delta Kappa Study Committee on Evaluation," American Educational Research Association New York, New York February 6, 1971

DETERMINING "MOST PROBABLE" CAUSES: A CALL FOR RE-EXAMINING EVALUATION METHODOLOGY

James L. Wardrop
University of Illinois at Urbana-Champaign

The major point of this presentation can be stated quite succinctly: namely, a central focus of educational evaluation is explanation (or, more precisely, the selection among several possible explanations). Once this thesis has been stated, the really hard work begins. I have now obligated myself to do three things: first, to explicate this succinct statement and attempt to give it substance; second, to justify the assertion I have made; and finally to indicate in some way how my thesis may be viewed as a reaction to the Phi Delta Kappa Study Commission materials. If I succeed in discharging any one of these obligations, my day will have been an unprecedented success, whether yours will or not.

A Notion of Causality

I circulated an earlier version of these comments to a number of friends, colleagues, and acquaintances (a few people actually fell into more than one of these three categories). A gratifying number of these people reacted. The reactions convinced me of two things: one, I had come up with the best projective technique for educational evaluators yet devised. (I will not take the time here to share with you the "projective" portions of those reactions.) Additionally, many of the reactions I received challenged my statement about the centrality of explanation to evaluation, for one reason. In that earlier draft, I made the following statement: "Explanation, as it is used in this paper, refers to the determination of the most probable cause for a phenomenon." Oh, the naiveté I exhibited in that sentence! I had forgotten that to use the word "cause" with people trained in the social sciences is much like sticking one's head into a beehive.

Nevertheless, I am going to stand by what I wrote then. (Naiveté dies hard in me!) One difference, though, is that I am going to try to clarify what I meant by "cause" in the context of this paper. Ernest Nagel, in his chapter on "Types of Causal Explanation in Science" in Lerner's book Cause and Effect, has considered what he referred to as "conditionally necessary causes." That is, suppose event E was observed. When E occurred, antecedent conditions A, B, and C were present. (It is possible, as Nagel pointed out, that we may be unaware of the existence of some or all of these conditions.) The general rule which applies to this situation might be stated as follows: Given that conditions A, B, and C are present, if condition D is also present, event E will occur; while if D is not present, E will not occur. Since condition D alone is not sufficient to cause the occurrence of E and since E may occur under some circumstances in the absence of D, we may speak of D as a contingently necessary cause of E. This is precisely the notion of causality I had in mind in writing that "explanation--the determination of the most probable cause or causes for a phenomenon--is a central focus of educational evaluation."

It is my contention that, in every type of evaluation presented by the PDK Commission, explanation is crucial. Further, I would argue that the PDK volume does not adequately treat this concept nor does it adequately consider some of the implications of the concept for the methodology of evaluation.

The Role of Explanation in Evaluation

In evaluation, as in experimentation, we seek to rule out, insofar as we are able, alternative explanations for phenomena. In context evaluation, we monitor the system in order to identify problems and isolate possible causes of these problems. Since the subsequent delineation of a class of possible change

strategies is directly dependent upon the causes so identified, it is vital that the evaluator be able to provide information of such quality as to insure that the identification of a cause or causes have a high probability of being correct. In other words, alternative explanations for the observed phenomenon (problem) must be shown to be unlikely.

In input evaluation, also, the issue is one of explanation, the attribution of causality. (If we do A, then X will be more likely to occur than if we do B or C. I.e., A is a more probable cause--as "cause" was defined earlier--of X than are B and C.) Once again, the decision (to do A, or B, or C) determines how and where and to what extent we are going to invest our resources. The ruling out of--or assignment of low probabilities to--alternate causal relationships is critical.

One major focus of process evaluation is upon the early identification and removal of barriers to the success of the particular program selected to implement the change strategy. As before, we are faced with the need for valid explanations. To call something a "barrier to success" is to make a causal inference of the form: if Q, then not X. That is, the occurrence of (existence of) Q reduces the likelihood that X will occur (increases the likelihood that "not X" will occur). Solving the problems of barriers is in this way formally equivalent to making the kinds of selection decisions which input evaluation serves, with the same implications relative to the attribution of causality.

Product evaluation can be thought of as representing the effort towards final verification of the web of explanations which has preceded it. If the causal relationships postulated earlier have been correct (if the explanations have been valid), then the hoped-for (intended) outcomes will occur. It is appropriate at this point to remind ourselves that other, unintended outcomes

will also occur. It is in connection with product evaluation that we most often bring to bear the wealth of inferential statistical methods, apply our principles of experimental design, and in general call up our methodological "big guns." The concern in the present paper is that we cannot afford to wait until this final stage to provide a sound methodological base for causal inference. The methodology of experimental design and traditional statistical techniques may not be--and probably are not--appropriate throughout the evaluation process (see below), but some methodologies must be employed which will provide us with a sound basis for our explanations.

The Search for Methodology

The preceding paragraphs have made a case for the centrality of "explanation" to evaluation as it is represented in the CIPP approach. On the basis of those arguments, one must agree that the ruling out of (or assigning low probabilities to) alternative explanations--or at least providing data upon which to base such decisions about alternative explanations--is an important aspect of evaluation.

While the distinction between research and evaluation is important and needs to be emphasized (as the PDK authors have done), I fear that a preoccupation with the differentiation may lead to an overly casual attitude on the part of some evaluators towards the quality of the information on which explanations produced within the evaluation setting are based. Threats to internal and--in some instances--external validity must receive extensive attention. If anything, they are even more important in an evaluation setting--where decisions (based on chains of causal inferences) determine the allocation of previous resources to a considerable degree--than in most research (especially basic

research) settings. If a researcher commits a Type I error, he (or other researchers) may pursue an inappropriate question until the error is discovered. The possible consequences of an evaluator (or decision-maker on the basis of information provided by the evaluator) committing the analogous kind of error are much more immediately felt in the resulting misallocation of resources.

The traditional model for educational research derives to a great extent from agricultural experimentation, after being filtered through experimental psychology. In his efforts to provide valid information on which to base explanations, the evaluator will often find this existing methodology both inadequate and inappropriate. In such circumstances, there are two alternatives to be considered. As a first step, and one which has considerable potential, we need to seek out from other disciplines--sociology, economics, anthropology, history, political science, etc.--methodologies for arriving at valid explanations. A second alternative, once inadequacies in methodology have been identified, is to set out to develop new approaches for gathering and analyzing information, in order to minimize the probability that alternative explanations are in fact correct.

Identifying Methodological Needs

In the preceding section, a task for evaluation methodologists was laid out. One essential aspect of that task is the identification of evaluation activities for which existing methodology is inadequate. Through an emphasis on the underlying search for causality, we should be able readily to identify many of those inadequacies. This approach leads directly to a concern for the nature of <u>evidence</u>. What kinds of evidence will best enable the evaluator (or decision-maker) to confidently discard alternative explanations as implausible?

How can the evidence the evaluator collects best be communicated to the decision-maker?

Given the position of the PDK Commission that evaluation serves the decision-maker, other very important questions arise: What kinds of evidence is the decision-maker willing to accept as bases for his inferences? Are these the kinds of evidence he should (according to some criteria) accept? The hope is that there is some commonality among decision-makers in terms of the kinds of evidence they are willing to accept, that the answer to this question does not depend entirely upon the idiosyncracies of the individual decision-makers, that given certain decision settings and decision types, decision-makers in common tend to seek certain kinds of evidence. Answering the "should" question will take much hard, logical thinking and--probably--years of investigation in an effort to validate the outcomes of that thinking.

Summary

If properly carried out, then, the task of the evaluator is in some ways much more difficult than that of the researcher. First, the evaluator finds himself working in naturalistic settings, settings in which many uncontrolled--and uncontrollable--sources of variation are operating. He is placed in the position of seeking consistent covariation over time and context, such covariation to be an important datum for his attempts at inferential explanation. Because the consequences of decisions based on evaluation data have considerable implication for (and effect on) the allocation of resources, it is imperative that gaps in extant evaluation methodology be identified and some of those resources allocated to closing the gaps.

You will have made an inference about my comments by now, one I wish to reinforce. (You have probably made several other inferences I would rather not reinforce, also.) Namely, I do not have any panaceas; I am not even sure where the answers will come from. But I expect to spend a considerable portion of my time in the near future worrying the issue of evidence, explanation, and causality in educational evaluation; and I hope others will do the same.